УДК 517.6

Theoretical and Numerical Result for Linear Optimization Problem Based on a New Kernel Function

Louiza Derbal*
Zakia Kebbiche†
Department of Mathematics, Faculty of Sciences
University of Ferhat Abbas
Setif1, 19000
Algeria

Received 09.07.2018, received in revised form 06.12.2018, accepted 16.01.2019

The propose of this paper is to improve the complexity results of primal-dual interior-point methods for linear optimization (LO) problem. We define a new proximity function for (LO) by a new kernel function wich is a combination of the classic kernel function and a barrier term. We present various proprieties of this new kernel function. Futhermore, we formilate an algorithm for a large-update primal-dual interior-point method (IPM) for (LO). It is shown that the iteration bound for large-update and smal-update primal-dual interior points methods based on this function is a good as the currently best know iteration bounds for these type of methods. This result decreases the gap between the practical behaviour of the large-update algorithms and their theoretical performance, which is an open problem.The primal-dual algorithm is implemented with different choices of the step size.

Numerical results show that the algorithm with practical and dynamic step sizes is more efficient than that with fixed (theoretical) step size.

Keywords: kernel function, interior point algorithms, linear optimization, complexity bound, primal-dual methods.
DOI: 10.17516/1997-1397-2019-12-2-160-172.

Introduction

In this paper we deal with interior point methods (IPMs) for linear optimization (LO). Since Karmarkar's seminal paper [5], many researchers have proposed and analyzed various IPMs for LO and a large amount of results have been reported. For a survey we refer to recent books on the subject [3, 8, 10, 12, 13]. In order to describe the idea of this paper we need to recall some ideas underlying new primal-dual IPMs. Recently, Peng, Roos and Terlaky [8] introduced the so-called self-regular barrier functions for primal-dual IPMs for LO and designed primal-dual interior-point algorithm based on self-regular proximities. Each such barrier function is determined by its (univariate) self-regular kernel function. The complexity bounds obtained by these authors are $O\Big(\sqrt{n}\log\frac{n}{\varepsilon}\Big)$ and $O\Big(\sqrt{n}\log n\log\frac{n}{\varepsilon}\Big)$, for small-update methods and large-update methods, respectively, which are currently the best known bounds. Motivated by their work, in this paper we present a new class of kernel functions which are not self-regular. The best iteration bound of large-update interior point methods based on these functions is shown to be $O\big(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon}\big)$ and for small-update methods is $O\big(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon}\big)$. These are currently the best-known bounds for primal-dual IPMs.

The paper is organized as follows. In Section 1, we start with some notations then we briefly review the basic concept of primal-dual IPMs for LO, such as the central path and new search

*l_derbal@yahoo.fr
†kebbichez@yahoo.fr

directions. The generic polynomial interior-point algorithm for LO is also presented. In Section 2, we define a new kernel function and present its properties. We analyze the algorithm and derive the complexity bound for large and small-update methods in Section 3. Numerical results are described in Section 4. Finally, Conclusion contains some conclusions and directions for future research.

1. Preliminaries

1.1. Notations

Some notations used throughout the paper are as follows. $\mathbb{R}^n$ is the n-dimensional Euclidean space with the inner product $\langle .,. \rangle$ and $\|.\|$ denotes the 2-norm. $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ denote the set of nonnegative vectors and the set of positive vectors, with n components, respectively. For x, $s \in \mathbb{R}^n$, $x_{\min}$ and xs denote the smallest component of the vector x and the componentwise product of vector x and s respectively. We denote by $X = diag(x)$ the $n \times n$ diagonal matrix with components of vector $x \in \mathbb{R}^n$ are the diagonal entries, e denotes the n-dimensional vector of ones. For $f,\ g : \mathbb{R}^n_{++} \to \mathbb{R}^n_{++}$, $f = O(g)$ if $f(x) \leqslant C_1 g(x)$ for some positive constant C_1 and $f = \theta(g)$ if $C_2 g(x) \leqslant f(x) \leqslant C_3 g(x)$ for some positive constants C_2 and C_3.

1.2. The central path

In this paper, we consider the linear optimization (LO) problem in standard form

$$\min \{\langle c, x\rangle : Ax = b,\ x \geqslant 0\}, \tag{P}$$

where $A \in \mathbb{R}^{m\times n}$ with $rank(A) = m$, $b \in R^m$ and $c \in \mathbb{R}^n$. The dual problem of (P) is given by

$$\max \left\{\langle b, y\rangle : A^T y + s = c,\ s \geqslant 0\right\}, \tag{D}$$

with $y \in \mathbb{R}^m$ and $s \in \mathbb{R}^n$. Without loss of generality [10] we assume that (P) and (D) satisfy the interior-point condition (IPC), i.e., there exist x^0, y^0 and s^0such that

$$Ax^0 = b,\ x^0 > 0, \qquad A^T y^0 + s^0 = c,\ s^0 > 0. \tag{1}$$

It is well known that finding an optimal solution of (P) and (D) is equivalent to solving the nonlinear system

$$Ax = b,\ x \geqslant 0, \qquad A^T y + s = c,\ s \geqslant 0, \quad xs = 0. \tag{2}$$

The basic idea of primal-dual IPMs is to replace the third equation in (2), the so-called complementarity condition for (P) and (D), by the parameterized equation $xs = \mu e$, with $\mu > 0$. Thus we consider the system

$$Ax = b,\ x \geqslant 0, \qquad A^T y + s = c,\ s \geqslant 0, \quad xs = \mu e. \tag{3}$$

Due to the last equation, any solution (x, y, s) of (3) will satisfy $x > 0$ and $s > 0$. Surprisingly enough, if the IPC is satisfied, then there exists a solution, for each $\mu > 0$, and this solution is unique. It is denoted as $(x(\mu), y(\mu), s(\mu))$, and we call $x(\mu)$ the μ-center of (P) and $(y(\mu), s(\mu))$ the μ-center of (D). The set of μ-centers is called the central path of (P) and (D). If $\mu \to 0$, then the limit of the central path exists, and since the limit points satisfy the complementarity condition, the limit yields optimal solutions for (P) and (D). IPMs follow the central path approximately. We briefly describe the usual approach. Without loss of generality, we assume that $(x(\mu), y(\mu), s(\mu))$ is known for some positive μ. We then decrease μ to $\mu := (1 - \theta)\mu$ for some fixed $\theta \in (0, 1)$, and we solve the following Newton system

$$A\Delta x = 0, \quad A^T \Delta y + \Delta s = 0, \quad x\Delta s + s\Delta x = \mu e - xs. \tag{4}$$

This process is repeated until μ is small enough, say until $n\mu \leqslant \varepsilon$; at this stage we have found an ε-solution of problems (P) and (D).

By taking a step along the search direction, one constructs a new triplet (x_+, y_+, s_+) with $x_+ = x + \alpha\Delta x,\ s_+ = s + \alpha\Delta s,\ y_+ = y + \alpha\Delta y$ where α denote the step size, $\alpha \in (0, 1)$, which has to be chosen appropriately (defined by some line search rules). If necessary, we repeat the procedure until we find iterates that are in a certain neighborhood of μ-center $(x(\mu), y(\mu), s(\mu))$.

1.3. Search directions

Now we introduce the scaled vector v and the scaled search directions d_x and d_s as follows

$$v = \sqrt{\frac{xs}{\mu}}, \quad d_x = \frac{v\Delta x}{x}, \quad d_s = \frac{v\Delta s}{s}. \tag{5}$$

System (4) can be rewritten as follows

$$\bar{A}d_x = 0, \quad \overline{A}^T\Delta y + d_s = 0, \quad d_x + d_s = v^{-1} - v. \tag{6}$$

Where $\bar{A} = \frac{1}{\mu}AV^{-1}X$, $V = diag(v)$ and $X = diag(x)$. Note that the right-hand side of the third equation in (6) equals to the negative gradient of the logarithmic barrier function $\Psi_l(v)$, i.e., $d_x + d_s = -\nabla\Psi_l(v)$. Where the barrier function $\Psi_l : \mathbb{R}^n_{++} \to \mathbb{R}_+$ is defined as follows

$$\Psi_l(v) = \sum_{i=1}^{n}\psi_l(v_i), \quad \psi_l(v_i) = \frac{v_i^2 - 1}{2} - \log v_i, \quad v_i > 0.$$

Note that $d_x = d_s = 0$ if and only if $v^{-1} - v = 0$ if and only if $x = x(\mu)$, $s = s(\mu)$. By replacing the proximity function $\Psi_l(v)$ by a proximity function $\Psi(v) = \sum_{i=1}^{n}\psi(v_i)$, where $\psi(t)$ is any strictly differentiable convex barrier function on R^n_{++}, with $\psi(1) = \psi'(1) = 0$, the system (6) is converted to the following system

$$\bar{A}d_x = 0, \quad \overline{A}^T\Delta y + d_s = 0, \quad d_x + d_s = -\nabla\Psi(v). \tag{7}$$

We reassert that in (7), $d_x = d_s = 0$ holds if and only if $\Psi(v) = 0$ if and only if $v = e$ if and only if $(x, s) = (x(\mu), s(\mu))$, as it should.

1.4. The generic interior-point algorithm for (LO)

Generic primal-dual IPMs for (LO)

Algorithm 1. *Input: a proximity function $\Psi(v)$; a threshold parameter $\tau > 1$; an accuracy parameter $\varepsilon > 0$; a fixed barrier update parameter $\theta, 0 < \theta < 1$;*

begin

$x := e; s := e; \mu := 1; v := e$

while $n\mu \geqslant \varepsilon$ do

begin (outer iteration)

$\mu := (1 - \theta)\mu;\ v := \dfrac{v}{\sqrt{1 - \theta}}$

while $\Psi(v) > \tau$ *do*
begin (inner iteration)
- *Find search directions by solving system* (7);
- *Determine a step size* α;
- *Put:* $x := x + \alpha\Delta x$; $s := s + \alpha\Delta$; $y := y + \alpha\Delta y$; $v := \sqrt{\frac{xs}{\mu}}$;

end (inner iteration),
end (outer iteration),
end.

Large and small-update methods

The parameters τ, θ and the step size α should be chosen in such a way that the algorithm is optimized in the sense that the number of iterations required by the algorithm is as small as possible. The choice of the so-called barrier update parameter θ plays an important role both in theory and practice of IPMs. Usually, if θ is a constant independent of the dimension n of the problem, then we call the algorithm a large-update (or long-step) method. If θ depends on the dimension of the problem, such as $\theta = O(n)$, then the algorithm is named a small-update (or short-step) method.

2. The new kernel function and its properties

In this paper, we define a new kernel function with logarithmic barrier term and propose primal-dual interior-point method which all the result of the complexity bound for large-update methods based on logarithmic kernel function, we prove that the corresponding algorithm has $O\Big(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon}\Big)$ complexity bound for large-update method and $O\Big(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon}\Big)$ for small-update method.

2.1. Properties of the new kernel function

We define a new function $\psi(t)$ as follows

$$\psi(t) = \frac{t^2 - 1 - \log(t)}{2} + \frac{e^{\frac{1}{tq}-1} - 1}{2q} \quad \text{for } t > 0, \quad q \geqslant 1. \tag{8}$$

Then, we have:

$$\begin{aligned}
&\psi'(t) = t - \frac{1}{2t} - \frac{e^{\frac{1}{tq}-1}}{2t^{q+1}},\\
&\psi''(t) = 1 + \frac{1}{2t^2} + \frac{1}{2}\Big(\frac{(q+1)t^q + q}{t^{2q+2}}\Big)e^{\frac{1}{tq}-1} > 1,\\
&\psi'''(t) = \frac{-1}{t^3} - \frac{1}{2}(q^2t^{-(3q+3)} + 3q(q+1)t^{-(2q+3)} + (q+1)(q+2)t^{-(q+3)})e^{\frac{1}{tq}-1} < 0.
\end{aligned} \tag{9}$$

We use $\Psi(v)$ as the proximity function to measure the distance between the current iterate and the μ-center for given $\mu > 0$. We also define the norm-based proximity measure, $\delta(v) : \mathbb{R}^n_{++} \to \mathbb{R}_+$, as follows

$$\delta(v) := \frac{1}{2}\left\|\nabla\Psi(v)\right\| = \frac{1}{2}\left\|d_x + d_s\right\|. \tag{10}$$

Lemma 1. *For* $\psi(t)$ *we have the following:* $\psi(t)$ *is exponentially convex for all* $t > 0$*; that is*

a) $\psi\left(\sqrt{t_1t_2}\right) \leqslant \frac{1}{2}\left(\psi(t_1) + \psi(t_2)\right)$,

b) $\psi''(t)$ *is monotonically decreasing for all* $t > 0$,
c) $t\psi''(t) - \psi'(t) > 0$ *for all* $t > 0$,
d) $\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0,\ t > 1,\ \beta > 1.$

Proof. For (a), using (9), we have $t\ \psi''(t) + \psi'(t) = 2t + \frac{1}{2}q(t^{-(q+1)} + t^{-(2q+1)}) > 0$ for all $t > 0$, and by Lemma 2.1.2 in [8], we have the result. For (b) and (c), using (9), so we have the result. For (d), using Lemma 2.4 in [2] , we have the result. This completes the proof. □

Lemma 2. *For* $\psi(t)$*, we have*

$$\begin{cases} \frac{1}{2}(t-1)^2 \leqslant \psi(t) \leqslant \frac{1}{2}(\psi'(t))^2, & t > 0, \quad (11) \\ \psi(t) \leqslant \frac{2+q}{2}(t-1)^2, & t > 1. \quad (12) \end{cases}$$

Proof. For (11) , since $\psi''(t) > 1$, we have $\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)d\zeta d\xi \leqslant \int_1^t \psi''(\xi)\psi'(\xi)d\xi = \frac{1}{2}(\psi'(t))^2$, and $\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta)d\zeta d\xi \geqslant \int_1^t \int_1^\xi 1 d\zeta d\xi = \frac{1}{2}(t-1)^2$. For (12), since $\psi(1) = \psi'(1) = 0$, $\psi'''(t) < 0$, $\psi''(1) = 2 + q$, and by using Taylor's theorem, we have $\psi(t) \leqslant \frac{2+q}{2}(t-1)^2$. This completes the proof. □

Lemma 3. *Let* $\varrho : [0, \infty) \to [1, \infty)$ *be the inverse function of* $\psi(t)$ *for* $t \geqslant 1$*. Then we have*

$$1 + \sqrt{\frac{2s}{q+2}} \leqslant \varrho(s) \leqslant 1 + \sqrt{2s}. \quad (13)$$

Proof. Let $s = \psi(t),\ t \geqslant 1,$ *i.e.,* $\varrho(s) = t,\ t \geqslant 1$. By the definition of $\psi(t)$ we have $s = \psi(t) \geqslant \geqslant \frac{1}{2}(t-1)^2$, which implies that $t = \varrho(s) \leqslant 1 + \sqrt{2s}$. By (12), we have $s = \psi(t) \leqslant \frac{2+q}{2}(t-1)^2,\ t > 1$, so $t = \varrho(s) \geqslant 1 + \sqrt{\frac{2s}{q+2}}$. □

In the next lemma we use the so-called barrier term $\psi_b(t)$ of $\psi(t)$, which is defined by $\psi(t) = \frac{t^2-1}{2} + \psi_b(t),\ t > 0$.

Lemma 4. *Let* $\rho : [0, \infty) \to (0, 1]$ *be the inverse function of the restriction of* $-\frac{\psi'(t)}{2}$ *in the interval* $(0, 1]$*,* $\underline{\rho} : [0, \infty) \to (0, 1]$ *be the inverse function of the restriction of* $-\psi_b'(t)$ *in the interval* $(0, 1]$ *and* $s_b = -\psi_b'(t)$*. Then one has*

$$\begin{cases} \rho(s) \geqslant \underline{\rho}(1 + 2s), & (14) \\ \underline{\rho}(s_b) \geqslant \frac{1}{(\log(2s_b)+1)^{\frac{1}{q}}}, \quad s_b > \frac{1}{2}. & (15) \end{cases}$$

Proof. Let $t = \rho(s)$. Due to definition of ρ as the inverse function of $-\frac{\psi'(t)}{2}$ for $t \leqslant 1$ this means that $-2s = \psi'(t) = t + \psi_b'(t),\ 0 < t \leqslant 1$. Since $t \leqslant 1$ this implies $-\psi_b'(t) = t + 2s \leqslant 1 + 2s$.

Function $\underline{\rho}(s_b)$ is also monotonically decreasing. We can say that $\rho(s) = t = \underline{\rho}(s_b) = = \underline{\rho}(-\psi_b'(t)) \geqslant \underline{\rho}(1 + 2s)$. For (15), let $s_b = \frac{1}{2\underline{\rho}(s_b)} + \frac{1}{2}\underline{\rho}(s_b)^{-(q+1)}e^{\underline{\rho}(s_b)^{-q}-1},\ 0 < \underline{\rho}(s_b) \leqslant 1$, $s_b > \frac{1}{2}$ means that $e^{\underline{\rho}(s_b)^{-q}-1} = 2\underline{\rho}(s_b)^{(q+1)}s_b - \underline{\rho}(s_b)^q \leqslant 2s_b$. Hence $\underline{\rho}(s_b) \geqslant \frac{1}{(\log(2s_b)+1)^{\frac{1}{q}}}$.

This completes the proof. □

Lemma 5. *Let* $\varrho : [0, +\infty) \to [1, +\infty)$ *be the inverse function of* $\psi(t), t \geqslant 1$. *Then we have* $\Psi(\beta v) \leqslant n\psi\Big(\beta\varrho\Big(\frac{\Psi(v)}{n}\Big)\Big)$, $v \in \mathbb{R}_{++}$, $\beta \geqslant 1$.

Proof. Using Lemma 1(d), and Theorem 3.2 in [1], we can get the result. This completes the proof. □

Lemma 6. *Let* $0 \leqslant \theta \leqslant 1, v_+ = \dfrac{v}{\sqrt{1-\theta}}$, *if* $\Psi(v) \leqslant \tau$, *then we have* $\Psi(v_+) \leqslant \dfrac{n\theta + 2\tau + 2\sqrt{2n\tau}}{2(1-\theta)}$.

Proof. Since $\dfrac{1}{\sqrt{1-\theta}} \geqslant 1$ and $\varrho\Big(\dfrac{\Psi(v)}{n}\Big) \geqslant 1$, then $\dfrac{\varrho\big(\frac{\Psi(v)}{n}\big)}{\sqrt{1-\theta}} \geqslant 1$, and for $t \geqslant 1$, we have $\psi(t) \leqslant$ $\leqslant \dfrac{t^2-1}{2}$. Using Lemma 5 with $\beta = \dfrac{1}{\sqrt{1-\theta}}$ (13) and $\Psi(v) \leqslant \tau$ we have $\Psi(v_+) \leqslant \dfrac{n\theta + 2\tau + 2\sqrt{2n\tau}}{2(1-\theta)}$. This completes the proof. □

Denote

$$\Psi_0 = \frac{n\theta + 2\tau + 2\sqrt{2n\tau}}{2(1-\theta)} = L(n, \theta, \tau) \tag{16}$$

then Ψ_0 is an upper bound for $\Psi(v)$ during the process of the algorithm.

3. Analysis of algorithm

In this section, we compute the feasible step size α such that the proximity function is decreasing and is bound for the decrease during inner iterations; then give the default step size $\tilde{\alpha}$; $\tilde{\alpha} = \dfrac{1}{1 + (2q+1)(1+4\delta)\left[\log(2+8\delta)+1\right]^{\frac{q+1}{q}}}$. We will show that the step size note only keeps the iterates feasible but also give rise to a sufficiently large decrease of the barrier function $\Psi(v)$ in each inner iteration. For fixed μ, taking a step size α, we have new iterates $x_+ := x + \alpha\Delta x$; $y_+ = y + \alpha\Delta y$; $s_+ := s + \alpha\Delta s$. Using (5), we have $x_+ = \dfrac{x}{v}(v + \alpha d_x)$; $s_+ = \dfrac{s}{v}(v + \alpha d_s)$; so we have $v_+ = \sqrt{\dfrac{x_+ s_+}{\mu}} = \sqrt{(v + \alpha d_x)(v + \alpha d_s)}$.

Define, for $\alpha > 0$, $f(\alpha) = \Psi(v_+) - \Psi(v)$. Then $f(\alpha)$ is the difference of proximity between a new iterate and a current iterate for fixed μ. By Lemma 1 (a), we have $\Psi(v_+) =$ $= \Psi(\sqrt{(v + \alpha d_x)(v + \alpha d_s)}) \leqslant \dfrac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s))$. Therefore, we have $f(\alpha) \leqslant f_1(\alpha)$, where

$$f_1(\alpha) = \frac{1}{2}(\Psi(v + \alpha d_x) + \Psi(v + \alpha d_s)) - \Psi(v). \tag{17}$$

Obviously, $f(0) = f_1(0) = 0$. Taking the first two derivatives of $f_1(\alpha)$ with respect to α, we have $f_1^{'}(\alpha) = \dfrac{1}{2}\sum\limits_{i=1}^{n}(\psi'(v_i + \alpha d_{xi})d_{xi} + \psi'(v_i + \alpha d_{si})d_{si})$, $f_1^{''}(\alpha) = \dfrac{1}{2}\sum\limits_{i=1}^{n}(\psi''(v_i + \alpha d_{xi})d_{xi}^2 + \psi''(v_i +$ $+\alpha d_{si})d_{si}^2)$. Using (7) and (10), we have $f_1^{'}(0) = \dfrac{1}{2}\nabla\Psi(v)^T(d_x + d_s) = -\dfrac{1}{2}\nabla\Psi(v)^T\nabla\Psi(v) =$ $= -2\delta(v)^2$. For convenience, we denote $v_1 = \min(v)$, $\delta := \delta(v)$.

Lemma 7. *Let* $\delta(v)$ *be as defined in (10). Then we have*

$$\delta(v) \geqslant \sqrt{\frac{1}{2}\Psi(v)}. \tag{18}$$

Proof. Using (11), we have $\Psi(v) = \sum_{i=1}^{n} \psi(v_i) \leqslant \sum_{i=1}^{n} \frac{1}{2}(\psi'(v_i))^2 = \frac{1}{2}\|\nabla\Psi(v)\|^2 = 2\delta(v)^2$, so $\delta(v) \geqslant \sqrt{\frac{1}{2}\Psi(v)}$. This completes the proof. □

Remark 1. *Throughout the paper, we assume that $\tau \geqslant 1$. Using Lemma 7 and the assumption that $\Psi(v) \geqslant \tau$, we have $\delta(v) \geqslant \sqrt{\frac{1}{2}}$.*

From Lemmas 4.1–4.3 in [2], we have the following Lemmas 8–11.

Lemma 8. *Let $f_1(\alpha)$ be as defined in (17) and $\delta(v)$ be as defined in (10). Then we have $f_1''(\alpha) \leqslant 2\delta^2\psi''(v_1 - 2\alpha\delta)$.*

Lemma 9. *$f_1'(\alpha) \leqslant 0$ certainly holds if α satisfies*

$$-\psi'(v_1 - 2\alpha\delta) + \psi'(v_1) \leqslant 2\delta. \tag{19}$$

Lemma 10. *Let $\rho : [0, +\infty) \to (0, 1]$ be the inverse function of $-\frac{1}{2}\psi'(t)$ for all $t \in (0, 1]$. Then the largest step size $\bar{\alpha}$ satisfying (19) is given by $\bar{\alpha} = \frac{1}{2\delta}(\rho(\delta) - \rho(2\delta))$.*

Lemma 11. *let ρ and $\bar{\alpha}$ be as defined in Lemma 10. then $\bar{\alpha} \geqslant \frac{1}{\psi''(\rho(2\delta))}$.*

Lemma 12. *let ρ and $\bar{\alpha}$ be as defined in Lemma 10. If $\Psi(v) \geqslant \tau \geqslant 1$, then we have*

$$\bar{\alpha} \geqslant \frac{1}{1 + (2q+1)(1+4\delta)\left[\log(2+8\delta) + 1\right]^{\frac{q+1}{q}}}.$$

Proof. Using Lemmas 11,4,7 and (9), we have

$$\bar{\alpha} \geqslant \frac{1}{\psi''(\rho(2\delta))} \geqslant \frac{1}{\psi''(\underline{\rho}(1+4\delta))}$$

by setting $t = \underline{\rho}(1+4\delta)$, $(0 < t \leqslant 1)$, it follows that

$$\begin{aligned}
\bar{\alpha} &\geqslant \frac{1}{\psi''(t)} = \frac{1}{1 + \frac{1}{2t^2} + \left[\frac{1}{2}(q+1)t^{-(q+2)} + \frac{1}{2}qt^{-(2q+2)}\right] e^{t^{-q}-1}} > \\
&> \frac{1}{1 + (2q+1)t^{-(q+1)}(-\psi_b'(t))} > \\
&> \frac{1}{1 + (2q+1)(1+4\delta)\left[\log(2+8\delta) + 1\right]^{\frac{q+1}{q}}}, \ \left(\text{put } t = \underline{\rho}\big(1 + \sqrt{\Psi(v)}\big)\right).
\end{aligned}$$

This completes the proof. □

Denoting

$$\tilde{\alpha} = \frac{1}{1 + (2q+1)(1+4\delta)\left[\log(2+8\delta) + 1\right]^{\frac{q+1}{q}}}, \tag{20}$$

we have that $\tilde{\alpha}$ is the default step size and that $\tilde{\alpha} \leqslant \bar{\alpha}$. From Lemma 1.3.3 in [8], we can get the following lemma

Lemma 13. *Suppose that $h(t)$ is a twice differentiable convex function with $h(0) = 0$, $h'(0) < 0$. Suppose that $h(t)$ attains its global minimum at $t^* > 0$ and $h''(t)$ is increasing with respect to t. Then, for any $t \in [0, t^*]$, we have $h(t) \leqslant \frac{th'(0)}{2}$.*

Let the univariate function h be such that $h(0) = f_1(0),\ h'(0) = f_1'(0) = -2\delta^2,\ \ h''(\alpha) = = 2\delta^2\psi''(v_1 - 2\alpha\delta)$. In this respect the next result is important.

Lemma 14. *Let $\tilde{\alpha}$ be the default step size as defined in (20) and let $\Psi(v) \geqslant 1$. Then*

$$f(\tilde{\alpha}) \leqslant -\frac{\sqrt{\Psi(v)}}{2 + (2q+1)(1+4\sqrt{2})\left[\log(2+4\sqrt{2\Psi_0}) + 1\right]^{\frac{q+1}{q}}}. \tag{21}$$

Proof. Using Lemma 4.5 in [2] and Remark 1, if the step size α satisfies $\alpha \leqslant \bar{\alpha}$, then $f(\alpha) \leqslant \leqslant -\alpha\delta^2$. So, for $\tilde{\alpha} \leqslant \bar{\alpha}$, we have $f(\tilde{\alpha}) \leqslant -\dfrac{\delta}{\sqrt{2} + (2q+1)(\sqrt{2}+4)\left[\log(2+8\delta) + 1\right]^{\frac{q+1}{q}}}$.

Since the decrease depends monotonically on δ, subtitution yields

$$f(\tilde{\alpha}) \leqslant -\frac{\sqrt{\Psi(v)}}{2 + (2q+1)(2+4\sqrt{2})\left[\log(2+4\sqrt{2\Psi_0})) + 1\right]^{\frac{q+1}{q}}},$$

where the last inequality follows from $\Psi_0 \geqslant \Psi \geqslant \tau \geqslant 1$. This result holds the lemma. □

3.1. Inner iteration bound

After the update of μ to $1-\mu$, we have $\Psi(v_+) \leqslant \dfrac{n\theta + 2\tau + 2\sqrt{2n\tau}}{2(1-\theta)} = L(n,\theta,\tau)$. We need to count how many inner iterations are required to return to the situation were $\Psi(v) \leqslant \tau$. we denote the value of $\Psi(v)$ after the μ update as Ψ_0; the subsequent values in the same outer iteration are denoted as Ψ_k , $k = 1, 2, \ldots, K$, where K denotes the total number of inner iterations in the outer iteration. The decrease in each inner iteration is given by (21). In [2] we can find the appropriate values of κ and $\gamma \in (0, 1]$: $\kappa = \dfrac{1}{2 + (2q+1)(2+4\sqrt{2})\left[\log(2+4\sqrt{2\Psi_0}) + 1\right]^{\frac{q+1}{q}}},\ \gamma = \dfrac{1}{2}$.

Lemma 15. *Let K be the number of inner iteration in the outer iteration. Then we have*

$$K \leqslant \left(4 + (2q+1)(4+8\sqrt{2})\left[\log(2+4\sqrt{2\Psi_0}) + 1\right]^{\frac{q+1}{q}}\right)\Psi_0^{\frac{1}{2}}.$$

Proof. By Lemma 1.3.2 in [8], we have

$$K \leqslant \frac{\Psi_0^{\gamma}}{\kappa\gamma} = \left(4 + (2q+1)(4+8\sqrt{2})\left[\log(2+4\sqrt{2\Psi_0}) + 1\right]^{\frac{q+1}{q}}\right)\Psi_0^{\frac{1}{2}}.$$

This completes the proof. □

3.2. Total iteration bound

The number of outer iterations is bounded above by $\dfrac{\log\frac{n}{\varepsilon}}{\theta}$ (see [10] Lemma II.17, page116). By multiplying the number of outer iterations by the number of inner iterations, we get an upper bound for the total number of iterations, namely,

$$\left(4 + (2q+1)(4+8\sqrt{2})\left[\log(2+4\sqrt{2\Psi_0}) + 1\right]^{\frac{q+1}{q}}\right)\Psi_0^{\frac{1}{2}}\frac{\log\frac{n}{\varepsilon}}{\theta}. \tag{22}$$

For large-update methods with $\tau = O(\sqrt{n})$ and $\theta = \Theta(1)$, we have $\Psi_0 = O(n)$ and $O(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon})$ iterations complexity.

Remark 2. *The better total iteration bounds is when $q = 1$, the total iteration bounds are $O\big(\sqrt{n}(\log\sqrt{n})^2\log\frac{n}{\varepsilon}\big)$ for large-update interior point methods.*

In case of a small-update methods, the best bound is obtained as follows. By (12), we have $\Psi(v_+) \leqslant n\psi\Big(\frac{1}{\sqrt{1-\theta}}\varrho\Big(\frac{\Psi(v)}{n}\Big)\Big) \leqslant \frac{n(2+q)}{2(1-\theta)}\left(\varrho(\frac{\Psi(v)}{n}) - \sqrt{1-\theta}\right)^2$. Using (13) and $\Psi(v) \leqslant \tau$ we have $\frac{n(2+q)}{2(1-\theta)}\left(\varrho(\frac{\Psi(v)}{n}) - \sqrt{1-\theta}\right)^2 \leqslant \frac{2+q}{2(1-\theta)}\left(\theta\sqrt{n} + \sqrt{2\tau}\right)^2 = \Psi_0$.

We have $\tau = O(1)$ and $\theta = \Theta\big(\frac{1}{\sqrt{n}}\big)$, in this case $\Psi_0 = O(q)$ and the iteration bound becomes $O\Big(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon}\Big)$ iteration complexsity.

4. Numerical results

The aim of this section is to investigate the influence of the choice of the new kernel function on the computational behavior of the generic primal-dual algorithm for linear optimizationas given in Fig. 1. The Algorithm is coded in MATLAB (R2014a) and our experiments were performed on PC with Processeur Genuine Intel(R) CPR T2080 @ 1,73GHZ installed memory (RAM) 2,00GO. For the parameters τ, θ and the accuracy parameter ε, we fixed these parameters to $\tau = \sqrt{n}, \theta \in \{0.3,\ 0.5,\ 0.7,\ 0.9,\ 0.99\}$ and $\varepsilon = 10^{-4}$.

The choice of the step size α $(0 < \alpha \leqslant 1)$ is another crucial issue in the analysis of the algorithm. It has to be made such that the closeness of the iterates to the current μ-center improves by a sufficient amount. In the theoretical analysis, the step size α is usually given a value that is very small during each inner iteration. In practice, this leads to very large inner iteration number. So, to accelerate the iteration process we propose a dynamic and practical choices defined bellow:

Dynamic choice [9]

We enlarge the step size by using the following procedure:

We take $\alpha = p\tilde{\alpha}$, when $p \geqslant 1$ is a fixed scalar according to the the size of the increment of x or s and $\tilde{\alpha}$ is the default step size (the theoretical choice).

In our numerical tests, for convenience, we set: $\alpha = \begin{cases} p_1\tilde{\alpha} & \text{if} \quad \|\Delta x\| \geqslant n \\ p_2\tilde{\alpha} & \text{if} \quad 1 \leqslant \|\Delta x\| \leqslant n \\ p_3\tilde{\alpha} & \text{if} \quad \|\Delta x\| \leqslant 1 \end{cases}$.

Practical choice [6]

We have the following conditions of strict positivity: $\begin{cases} x + \alpha_x^+ dx > 0 \\ s + \alpha_s^+ ds > 0 \end{cases}$.

Which give: $\alpha_x^+ = \beta\alpha_x$ and $\alpha_s^+ = \beta\alpha_s$, such as $0 < \beta < 1$, or

$\alpha_x = \begin{cases} \min\Big(-\frac{x_i}{dx_i}\Big) \text{ with } i \in I = \{i\colon dx_i < 0\} \\ 1 \qquad\qquad\qquad\qquad \text{elsewhere} \end{cases}$; $\alpha_s = \begin{cases} \min\Big(-\frac{s_i}{ds_i}\Big) \text{ with } i \in I = \{i\colon ds_i < 0\} \\ 1 \qquad\qquad\qquad\qquad \text{elsewhere} \end{cases}$.

We take $\alpha_k = \min(\alpha_x, \alpha_s)$. So the new iterated is $(x^+, s^+) = (x, s) + \alpha_k(dx, ds)$.

Example 1. We consider a linear program with $m = 5, n = 9$,

$$A = \begin{pmatrix} 0 & 1 & 2 & -1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 & -1 & 0 & 1 & 0 & 0 \\ -1 & 0 & -2 & 1 & 2 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & -1 & -2 & 0 & 0 & 0 & 1 \\ 1 & 3 & 4 & 2 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad c = \begin{pmatrix} 1 & 0 & -2 & 1 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}^T$$

and $b = \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \end{pmatrix}^T$. The starting point is

$x_0 = [\ 0,1819 \quad 0,0699 \quad 0,063 \quad 0,1105 \quad 0,2012 \quad 0,6732 \quad 1,1885 \quad 2,835 \quad 2,1912\ ]^T$,
$s_0 = [\ 4,939 \quad 3,544 \quad 4,7186 \quad 9,1788 \quad 4,5072 \quad 1,384 \quad 0,875 \quad 0,4241 \quad 0,4463\ ]^T$,
$y_0 = [\ -1,3843 \quad -0,8751 \quad -0,4241 \quad -0,4463 \quad -3,0424\ ]^T$.

The optimal solution is:

$x^* = (\ 0 \quad 0 \quad 0,2664 \quad 0 \quad 0 \quad 0,4269 \quad 1,1406 \quad 3,5729 \quad 2\ )^T$,
$y^* = \begin{pmatrix} 0 & 0 & 0 & 0 & -0,4999 \end{pmatrix}^T$,
$s^* = \begin{pmatrix} 1,5 & 1,4999 & 0 & 1,9999 & 1,4999 & 0 & 0 & 0 & 0 \end{pmatrix}^T$.

In the tables of results, n represents the size of the example, (Outer) represents the number of outer iterations, (Inner) represents the number of inner iterations and (Time) represents the calculation time in seconds.

Tab. 1 gives the numbers of inner and outer iterations for Example 1 with fixed scalars $p_1 = 100$, $p_2 = 50$ and $p_3 = 25$. We obtain the following results:

Table 1. Numbers of inner and outer iterations for Example with a fixed scalar p.

Step size choices	Inner	Outer	Time
Theoretical choice	2704	5	$24.627917s$
Dynamic choice	23	5	$0.218753s$
Practical choice	4	5	$0.109279s$

Example 2. We consider a linear program with $m = 3$, $n = 6$,

$$A = \begin{pmatrix} 2 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \quad c = \begin{pmatrix} 3 & -1 & 1 & 0 & 0 & 0 \end{pmatrix}^T \text{ and } b = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T.$$

The starting point is

$x_0 = [\,0.06757,\ 0.13258,\ 0.13302,\ 0.26774,\ 0.13302,\ 0.2664\,]^T$,
$s_0 = [1, 0, 4, 6, 1, 5, 1]^T$,
$y_0 = [-2, -2, -3]^T$.

The optimal solution is:

$x^* = (0.0000, 0.5000, 0.0000, 0.5000, 0.0000, 0.0004)^T$,
$y^* = (-0.5000, -0.4902, -0.5000)^T$,
$s^* = (4.5000, 0.0000, 1.9902, 0.0000, 0.9902, 0.0098)^T$.

There is a parameter p involved in the definition of the dynamic choice, we used several values of this parameter as indicated in Tabs. 2, 3 below. Theses values were chosen after some preliminary experiments that showed that these values gave the most promising iteration counts.

The following table gives the numbers of iterations for possible combinations of θ and p. The value $\theta = 0,9$ gives the lowest iteration count in all cases.

Table 2. Numbers of inner and outer iterations for several choices of θ and p

θ	p_1	p_2	p_3	Inner	Outer	Time
0.3	200	100	50	1	31	0.027893
0.5	201	100	50	1	16	0.029697
0.7	201	100	50	2	10	0.032418
0.9	423	100	50	2	5	0.019141
0.99	422	100	50	13	3	0.074557

Tab. 3 gives the numbers of inner and outer iterations for Example 2 with $\theta = 0.9$ and variable values of scalars p_1, p_2 and p_3.

Table 3. Numbers of inner and outer iterations for variable values of the scalar p

Step size choices	Inner	Outer	Time
Theoretical choice	2174	5	$7.715739s$
Dynamic choice	$p_1 = 100 \quad p_2 = 50 \quad p_3 = 25$ 21	5	$0.1593165s$
Dynamic choice	$p_1 = 423 \quad p_2 = 100 \quad p_3 = 50$ 2	5	$0.019141s$
Practical choice	4	5	$0.024999s$

Example 3. We consider the following example: $n = 2m$.

$$A(i,j) = \begin{cases} 0 \text{ if } i \neq j \text{ and } i \neq j+m \\ 1 \text{ if } i = j \text{ and } i = j+m, \end{cases} \quad c(i) = -1,\ c(i+m) = 0,\ b(i) = 2 \text{ for } i = 1, \ldots, m.$$

The starting point is: $x_0(i) = x_1(i+m) = 1$, $s_0(i) = 1$, $s_0(i+m) = 2$ and $y_0(i) = -2$ for $i = 1, \ldots, m$. The optimal solutions are obtained as follows:

$$x^* = \begin{cases} 2 \text{ if } i = 1, \ldots, m \\ 0 \text{ if } i = m+1, \ldots, n, \end{cases} \quad y^* = -1 \text{ for } i = 1, \ldots, n \text{ and } s^* = \begin{cases} 0 \text{ if } i = 1, \ldots, m \\ 1 \text{ if } i = m+1, \ldots, n. \end{cases}$$

We have the results with $\theta = 0.9$ in Tab. 4.

Table 4. Numbers of inner and outer iterations for several choices of the step size α for an example with variable size

n	p			Theoretical choice			Dynamic choice			Practical choice		
	p_1	p_2	p_3	Inner	Outer	Time	Inner	Outer	Time	Inner	Outer	Time
20	500	350	150	4171	6	$484.938590s$	4	6	$0.208902s$	4	6	$0.122828s$
50	1050	350	150	6977	6	$791.156169s$	3	6	$0.491385s$	4	6	$0.279962s$
100	1050	350	150	10385	6	$3475.590158s$	4	6	$2.937549s$	4	6	$0.360571s$
200	2000	350	280	15547	7	$19737.654690s$	5	7	$12.924049s$	5	7	$4.378243s$
400	3010	500	280	...	...	...	4	7	$68.116663s$	5	7	$21.679448s$
500	3110	510	280	...	...	...	5	7	$137.160435s$	4	7	$42.884622s$
1000	5525	510	350	...	...	...	6	7	1321.177117	5	7	$262.051496s$

Conclusion

In this paper, we have proposed a primal-dual interior point algorithm for (LO) based on a new kernel function. For this parametric kernel function, we have shown that the best result of iteration bounds for large and small-update methods can be achieved, namely $O(q\sqrt{n}(\log\sqrt{n})^{\frac{q+1}{q}}\log\frac{n}{\varepsilon})$ for large-update and $O(q^{\frac{3}{2}}(\log\sqrt{q})^{\frac{q+1}{q}}\sqrt{n}\log\frac{n}{\varepsilon})$ for small-update methods. In practice, the step size α plays a crucial role in the computational behavior of the algorithm. To accelerate the iteration process of our algorithm, we have proposed a dynamic and practical choices. The algorithm with practical step size work faster than that with the dynamic one, but for suitable values of the parameter p the two choices lead to a significant decrease in the total number of iterations.

For further research, it is necessary to think of a simple strategy to determine the appropriate values of the parameter p which keeps the iteration in the interior of the feasible domain. Furthermore, this algorithm may be possible extended to the semidefinite linear optimization, quadratic programming and linear complementarity problem with these choices of the step size.

References

[1] M.Achache, A new parameterized kernel function for LO yielding the best known iteration bound for a large-update interior point algorithm, *Afrika Matematika*, **27**(2016), no. 3–4, 591–601.

[2] Y.Q.Bai, M.El.Ghami, C.Roos, A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization, *SIAM Journal on Optimization*, **15**(2004), no. 1, 101–128.

[3] D.Den Hertog, Interior point approach to linear, quadratic, and convex programming, in: Mathematics and its Applications, vol. 277, Kluwer Academic Publishers, Dordrecht, 1994.

[4] M.Bouafia, D.Benterki, A.Yassine, An efficient parameterized logarithmic kernel function for linear optimization, Optimization Letters, 2018.

[5] N.K.Karmarkar, A new polynomial-time algorithm for linear programming, in: Proceedings of the 16th Annual ACM Symposium on Theory of Computing, 1984, 302–311.

[6] A.Keraghel, Etude adaptative et comparative des principales variantes dans l'algorithme de karmarkar, Thèse de Doctorat, Université de Joseph Fourier, Grenoble I, France, 1989.

[7] B.Kheirfam, M.Haghighi, A full-Newton step infeasible interior-point method for linear optimization based on a trigonometric kernel function, *Journal of Mathematical Programming and Operations Research*, **65**(2016), no. 4, 841–857.

[8] J.Peng, C.Roos, T.Terlaky, Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms, Princeton University Press, 2002.

[9] Z.G.Qian, Y.Q.Bay, Primal-dual Interior-Point Algorithms with Dynamic Step SizeBased on kernel functions for linear Programming, Journal of Shanghai University, **9**(2005), no. 5, 391–396.

[10] C.Roos, T.Terlaky, J.-Ph.Vial, Theory and Algorithms for Linear Optimization, An Interior-Point Approach, John Wiley & Sons, Chichester, UK, 1997.

[11] G.Sonnevend, An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming, in: Prekopa, A, J. Szelezsan, B. (Eds.), System Modelling and Optimization: Proceedings of 12th IFIP-Conference, Budapest, Hungary, 1985; Lecture Notes in Control and Inform. Sci., Springer, Berlin, vol. 84, 1986, 866–876.

[12] R.J.Vanderbei, Linear Programming, Foundations and Extensions, 2nd ed., in: International Series in Operations Research and Management Science, vol. 7, 1997.

[13] Y.Ye, Interior Point Algorithms, Theory and Analysis, John Wiley and Sons, Chichester, UK, 1997.

Теоретический и численный результат для задачи линейной оптимизации на основе новой функции ядра

Луиза Дербал
Закия Кеббиш
Кафедра математики, факультет наук
Университет Ферхата Аббаса
Сетиф-1, 19000
Алжир

Целью данной работы является улучшение результатов сложности первично-двойственных методов внутренней точки для задачи линейной оптимизации (LO). Мы определим новую функцию близости для (LO) новой функцией ядра, которая является комбинацией классической функции ядра и барьерного члена. Мы представляем различные свойства этой новой функции ядра. Кроме того, мы сформулируем алгоритм для большого обновления метода первичной-двойной внутренней точки (IPM) для (LO). Показано, что оценка итераций для методов простого обновления и малых обновлений, основанных на этой функции, наилучшая из известных в настоящее время границ итераций для методов этого типа. Этот результат уменьшает разрыв между практическим поведением алгоритмов с большим обновлением и их теоретической эффективностью, что является открытой проблемой. Алгоритм первичного двойственного типа реализован с различными вариантами выбора размера шага.

Численные результаты показывают, что алгоритм с практическим и динамическим размером шага более эффективен, чем алгоритм с фиксированным (теоретическим) размером шага.

Ключевые слова: функция ядра, алгоритмы внутренних точек, линейная оптимизация, оценка сложности, примало-дуальные методы.